

DeAcero - Predicción de Precios de Varilla

1. Introducción y Contexto

Como parte del proceso de selección para el puesto de Científico de datos en DeAcero, debe desarrollar una solución técnica que demuestre sus habilidades en ingeniería de datos, modelado predictivo y despliegue de soluciones.

DeAcero, como una de las siderúrgicas líderes en México, requiere capacidades predictivas para optimizar sus decisiones de compra de materia prima y estrategias de precio. El precio de la varilla corrugada es un indicador crítico que impacta directamente los márgenes operativos de la empresa.

2. Objetivo de la Prueba

Desarrollar y desplegar un API REST que prediga el precio de cierre del día siguiente para la varilla corrugada, utilizando datos históricos disponibles públicamente.

3. Requerimientos Técnicos

3.1 Endpoint Principal

Debe exponer UN único endpoint público accesible por internet:

GET /predict/steel-rebar-price

Respuesta esperada (formato JSON):

```
{  
  "prediction_date": "2025-01-XX",  
  "predicted price usd_per_ton": 750.45,  
  "currency": "USD",  
  "unit": "metric ton",  
  "model_confidence": 0.85,  
  "timestamp": "2025-01-XX T00:00:00Z"
```

3.2 Fuentes de Datos Sugeridas (No Obligatorias)

- London Metal Exchange (LME) - Precios de metales
- Trading Economics - Indicadores de commodities
- FRED (Federal Reserve Economic Data) - Series económicas
- World Bank Commodity Price Data
- Quandl/Nasdaq Data Link
- Yahoo Finance (para proxies como futuros de acero)

Nota: Puede utilizar cualquier fuente de datos públicos que considere relevante. La calidad y relevancia de los datos seleccionados será parte de la evaluación.

3.3 Restricciones y Consideraciones

1. Autenticación: El endpoint debe requerir un header de autorización:
2. X-API-Key: [valor_que_usted_defina]
3. Rate Limiting: Implemente un limite de 100 requests por hora por API key.
4. Cache: Las predicciones deben tener un cache de máximo 1 hora (para evitar recálculos innecesarios).
5. Documentación Mínima: Incluya en la respuesta del endpoint raiz (GET) la siguiente información:

```
{  
  "service": "Steel Rebar Price Predictor",  
  "version": "1.0",  
  "documentation url": "[URL a su documentación]",  
  "data sources": ["lista de fuentes utilizadas"],  
  "last model update": "timestamp"  
}
```

4. Criterios de Evaluación

4.1 Evaluación Cuantitativa (60% del puntaje)

- Se realizarán llamadas al API durante 5 días consecutivos
- Se comparará la predicción con el precio real del día siguiente
- Métrica: MAPE (Mean Absolute Percentage Error)
- Criterio de eliminación: El 50% de candidatos con mayor error promedio quedará excluido

4.2 Evaluación Cualitativa (40% del puntaje)

- Ingeniería de Features (15%): Creatividad y relevancia de las variables utilizadas
- Robustez del Sistema (10%): Manejo de errores, disponibilidad del servicio
- Calidad del Código (10%) : Estructura, documentación, mejores prácticas
- Escalabilidad (5%): Diseño que permita crecimiento futuro

5. Entregables

1. URL del Endpoint: Dirección pública accesible del API

2. API Key: Para acceder al servicio

3. Repositorio de Código: Link a GitHub/GitLab con:

- Código fuente
- README con instrucciones de despliegue
- Descripción del modelo y features utilizados
- Justificación de decisiones técnicas

6. Consideraciones Adicionales (Opcionales pero Valoradas)

- Monitoreo: Dashboard o métricas de performance del modelo
- A/B Testing: Capacidad de probar múltiples modelos
- Explicabilidad: Endpoint adicional que explique los factores que más influyen en la predicción
- Datos Complementarios: Uso de indicadores económicos, tipos de cambio, índices industriales

7. Restricciones y Lineamientos

- Presupuesto Cloud: La solución debe poder operar con menos de \$5 USD/mes
- Lenguajes Permitidos: Python, R, Java, Node.js, Go
- Tiempo de Respuesta: El endpoint debe responder en menos de 2 segundos
- Sin Dependencias Comerciales: No utilizar APIs de pago o servicios que requieran licencias

8. Preguntas Frecuentes

P: ¿Puedo usar modelos pre-entrenados? R: Si, pero debe documentar claramente qué utilizó y cómo lo adaptó al problema.

P: ¿Qué pasa si no hay datos de precio para un día específico (fin de semana, feriados)?

R: Su modelo debe manejar estos casos. Documente su estrategia.

P: ¿Puedo incluir endpoints adicionales? R: Si, pero solo el endpoint principal será evaluado para la métrica de error.

9. Plazo de Entrega

- Fecha límite: 7 días calendario desde la recepción de estas instrucciones
- Inicio de evaluación: Al día siguiente de la entrega
- Período de evaluación: 5 días consecutivos

10. Tips y Sugerencias (Sin dar la solución)

- Considere que los precios de commodities tienen patrones estacionales y tendencias
- Los eventos geopolíticos pueden causar volatilidad considere incluir algún indicador de incertidumbre
- La varilla corrugada está correlacionada con el precio del mineral de hierro y el carbón de coque
- Los tipos de cambio pueden influir en los precios locales vs internacionales
- Un modelo simple bien implementado es mejor que uno complejo mal ejecutado